1 GCCGTCACTCCCTCCGTCATCGATAACATCCTGTCCAAGATCGAGAACGAGTACG
1▶ AlaValThrProSerValIleAspAsnIleLeuSerLysIleGluAsnGluTyrG

56 AGGTGCTGTACCTGAAGCCCCTGGCAGGGGTCTACCGGAGCCTGAAGAAGCAG
19▶ luValLeuTyrLeuLysProLeuAlaGlyValTyrArgSerLeuLysLysGln

109 CTGGAGAACAACGTGATGACCTTCAACGTGAACGTGAAGGATATCCTGAACAGC
37▶ LeuGluAsnAsnValMetThrPheAsnValAsnValLysAspIleLeuAsnSer

163 CGGTTCAACAAGCGGGAGAACTTCAAGAACGTGCTGGAGAGCGATCTGATCCC
55▶ ArgPheAsnLysArgGluAsnPheLysAsnValLeuGluSerAspLeuIlePr

216 CTACAAGGATCTGACCAGCAGCAACTACGTGGTCAAGGATCCCTACAAGTTCC
72▶ oTyrLysAspLeuThrSerSerAsnTyrValValLysAspProTyrLysPheL

269 TGAACAAGGAGAAGAGAGATAAGTTCCTGAGCAGTTACAACTACATCAAGGATAG
90▶ euAsnLysGluLysArgAspLysPheLeuSerSerTyrAsnTyrIleLysAspSe

324 CATTGATACCGATATCAACTTCGCCAACGATGTCCTGGGATACTACAAGATCCT
108▶ rIleAspThrAspIleAsnPheAlaAsnAspValLeuGlyTyrTyrLysIleLe

378 GTCCGAGAAGTACAAGAGCGATCTGGATTCAATCAAGAAGTACATCAACGATAA
126▶ uSerGluLysTyrLysSerAspLeuAspSerIleLysLysTyrIleAsnAspLy

432 GCAGGGAGAGAACGAGAAGTACCTGCCCTTCCTGAACAACATCGAGACCCTGTA
144▶ sGlnGlyGluAsnGluLysTyrLeuProPheLeuAsnAsnIleGluThrLeuTy

486 CAAGACCGTCAACGATAAGATTGATCTGTTCGTGATCCACCTGGAGGCCAAGGT
162▶ rLysThrValAsnAspLysIleAspLeuPheValIleHisLeuGluAlaLysVa

NdeI

540 CCTGAACTACACATATGAGAAGAGCAACGTGGAGGTCAAGATCAAGGAGCTGAA
180▶ lLeuAsnTyrThrTyrGluLysSerAsnValGluValLysIleLysGluLeuAs

594 TTACCTGAAGACCATCCAGGATAAGCTGGCCGATTTCAAGAAGAACAACAACTT
198▶ nTyrLeuLysThrIleGlnAspLysLeuAlaAspPheLysLysAsnAsnAsnPh

648 CGTCGGGATCGCCGATCTGAGCACCGATTACAACCACAACAACCTGCTGACCAA
216▶ eValGlyIleAlaAspLeuSerThrAspTyrAsnHisAsnAsnLeuLeuThrLy

702 GTTCCTGAGCACCGGTATGGTCTTCGAAAACCTGGCCAAGACCGTCCTGAGCAA
234▶ sPheLeuSerThrGlyMetValPheGluAsnLeuAlaLysThrValLeuSerAs

756 CCTGCTGGATGGAAACCTGCAGGGGATGCTGAACATCAGCCAGCACCAGTGTGT
252▶ nLeuLeuAspGlyAsnLeuGlnGlyMetLeuAsnIleSerGlnHisGlnCysVa

810 GAAGAAGCAGTGTCCCCAGAACAGCGGGTGTTTCAGACACCTGGATGAGAGAGA
270▶ lLysLysGlnCysProGlnAsnSerGlyCysPheArgHisLeuAspGluArgGl

864 GGAGTGTAAGTGTCTGCTGAACTACAAGCAGGAAGGTGATAAGTGTGTGGAAAAC
288▶ uGluCysLysCysLeuLeuAsnTyrLysGlnGluGlyAspLysCysValGluAsn

919 CCCAATCCTACTTGTAACGAGAACAATGGTGGATGTGATGCCGATGCCAAGTGTACCG
307▶ ProAsnProThrCysAsnGluAsnAsnGlyGlyCysAspAlaAspAlaLysCysThrG

977 AGGAGGATTCAGGGAGCAACGGGAAGAAGATCACCTGTGAGTGTACCAAGCCTGATT
326▶ luGluAspSerGlySerAsnGlyLysLysIleThrCysGluCysThrLysProAspS

1034 CTTATCCACTGTTCGATGGTATCTTCTGTAGT
345▶ erTyrProLeuPheAspGlyIlePheCysSer

FIG. 1

```
1     GCAGTAACTCCTTCCGTAATTGATAACATACTTTCTAAAATTGAAAATGAATA
1▶    AlaValThrProSerValIleAspAsnIleLeuSerLysIleGluAsnGluTyrG
                         EcoNI (73)
56    AGGTTTTATATTTAAAACCTTTAGCAGGTGTTTATAGAAGTTTAAAAAAACAATT
19▶   luValLeuTyrLeuLysProLeuAlaGlyValTyrArgSerLeuLysLysGlnLe
111   AGAAAAATAACGTTATGACATTTAATGTTAATGTTAAGGATATTTTAAATTCACGA
37▶   uGluAsnAsnValMetThrPheAsnValAsnValLysAspIleLeuAsnSerArg
166   TTTAATAAACGTGAAAATTTCAAAAATGTTTTAGAATCAGATTTAATTCCATATA
56▶   PheAsnLysArgGluAsnPheLysAsnValLeuGluSerAspLeuIleProTyrL
221   AAGATTTAACATCAAGTAATTATGTTGTCAAAGATCCATATAAATTTCTTAATAA
74▶   ysAspLeuThrSerSerAsnTyrValValLysAspProTyrLysPheLeuAsnLy
276   AGAAAAAAGAGATAAATTCTTAAGCAGTTATAATTATATTAAGGATTCAATAGAT
92▶   sGluLysArgAspLysPheLeuSerSerTyrAsnTyrIleLysAspSerIleAsp
331   ACGGATATAAATTTTGCAAATGATGTTCTTGGATATTATAAAATATTATCCGAAA
111▶  ThrAspIleAsnPheAlaAsnAspValLeuGlyTyrTyrLysIleLeuSerGluL
386   AATATAAATCAGATTTAGATTCAATTAAAAAATATATCAACGACAAACAAGGTGA
129▶  ysTyrLysSerAspLeuAspSerIleLysLysTyrIleAsnAspLysGlnGlyGl
441   AAATGAGAAATACCTTCCCTTTTTAAACAATATTGAGACCTTATATAAAACAGTT
147▶  uAsnGluLysTyrLeuProPheLeuAsnAsnIleGluThrLeuTyrLysThrVal
496   AATGATAAAATTGATTTATTTGTAATTCATTTAGAAGCAAAAGTTCTAAATTATA
166▶  AsnAspLysIleAspLeuPheValIleHisLeuGluAlaLysValLeuAsnTyrT
551   CATATGAGAAATCAAACGTAGAAGTTAAAATAAAAGAACTTAATTACTTAAAAAC
184▶  hrTyrGluLysSerAsnValGluValLysIleLysGluLeuAsnTyrLeuLysTh
606   AATTCAAGACAAATTGGCAGATTTTAAAAAAAATAACAATTTCGTTGGAATTGCT
202▶  rIleGlnAspLysLeuAlaAspPheLysLysAsnAsnAsnPheValGlyIleAla
661   GATTTATCAACAGATTATAACCATAATAACTTATTGACAAAGTTCCTTAGTACAG
221▶  AspLeuSerThrAspTyrAsnHisAsnAsnLeuLeuThrLysPheLeuSerThrG
716   GTATGGTTTTTGAAAATCTTGCTAAAACCGTTTTATCTAATTTACTTGATGGAAA
239▶  lyMetValPheGluAsnLeuAlaLysThrValLeuSerAsnLeuLeuAspGlyAs
771   CTTGCAAGGTATGTTAAACATTTCACAACACCAATGCGTAAAAAAACAATGTCCA
257▶  nLeuGlnGlyMetLeuAsnIleSerGlnHisGlnCysValLysLysGlnCysPro
826   CAAAATTCTGGATGTTTCAGACATTTAGATGAAAGAGAAGAATGTAAATGTTTAT
276▶  GlnAsnSerGlyCysPheArgHisLeuAspGluArgGluGluCysLysCysLeuL
881   TAAATTACAAACAAGAAGGTGATAAATGTGTTGAAAATCCAAATCCTACTTGTAA
294▶  euAsnTyrLysGlnGluGlyAspLysCysValGluAsnProAsnProThrCysAs
936   CGAAAATAATGGTGGATGTGATGCAGATGCCAAATGTACCGAAGAAGATTCAGGT
312▶  nGluAsnAsnGlyGlyCysAspAlaAspAlaLysCysThrGluGluAspSerGly
991   AGCAACGGAAAGAAAATCACATGTGAATGTACTAAACCTGATTCTTATCCACTTT
331▶  SerAsnGlyLysLysIleThrCysGluCysThrLysProAspSerTyrProLeuP
                     PstI   (1059)
1046  TCGATGGTATTTTCTGCAGTCACCACCACCACCACCACTAAACT
349▶  heAspGlyIlePheCysSerHisHisHisHisHisHis•••
```

FIG. 2

Codon	AA	goat b-casei	goat K-casein	MSP wt	Edited MSP	mouse b-casei	mouse a-casei	mouse g-casein	mouse e-casei
TTT	Phe	5	4	8	0	4	8	3	4
TTC	Phe	4	3	7	15	4	6	7	1
TTA	Leu	0	2	25	0	0	0	0	0
TTG	Leu	0	2	3	0	0	0	0	1
TCT	Ser	5	1	4	1	13	5	7	5
TCC	Ser	2	2	2	3	6	14	8	2
TCA	Ser	1	4	10	1	1	3	2	0
TCG	Ser	0	1	0	0	0	0	0	0
TAT	Tyr	2	7	17	2	1	3	2	1
TAC	Tyr	1	2	3	18	2	6	6	7
TAA	***	1	2	0	0	1	0	1	0
TAG	***	0	0	0	0	0	0	0	0
TGT	Cys	1	1	10	12	0	0	1	0
TGC	Cys	0	2	2	0	2	2	2	1
TGA	***	0	0	0	0	0	1	0	1
TGG	Trp	1	1	0	0	0	2	2	2
CTT	Leu	9	1	9	0	16	9	3	3
CTC	Leu	5	2	0	0	7	8	0	1
CTA	Leu	1	2	1	0	1	2	1	0
CTG	Leu	11	5	0	38	10	17	4	1
CCT	Pro	17	6	4	2	8	6	3	0
CCC	Pro	12	0	1	6	8	6	6	4
CCA	Pro	3	13	5	1	5	6	2	2
CCG	Pro	1	1	0	1	0	0	0	1
CAT	His	0	1	3	0	2	6	2	1
CAC	His	5	3	1	4	4	0	3	0
CAA	Gln	5	9	9	0	9	21	9	7
CAG	Gln	16	6	0	9	21	32	12	8
CGT	Arg	0	1	1	0	0	0	0	0
CGC	Arg	0	0	0	0	1	0	0	0
CGA	Arg	0	0	1	0	0	0	0	1
CGG	Arg	1	0	0	3	0	0	0	0
ATT	Ile	4	5	13	0	3	4	3	4
ATC	Ile	6	3	2	20	7	5	8	5
ATA	Ile	1	3	5	0	1	0	2	0
ATG	Met	7	3	3	3	4	12	2	13
ACT	Thr	7	6	3	2	6	5	1	4
ACC	Thr	2	7	3	13	4	4	4	4
ACA	Thr	2	4	9	1	1	1	2	0
ACG	Thr	0	0	1	0	0	0	2	0
AAT	Asn	2	6	29	3	4	6	3	1
AAC	Asn	2	3	12	38	4	9	4	6
AAA	Lys	7	6	38	0	6	7	3	5
AAG	Lys	6	4	4	42	3	6	13	7
AGT	Ser	2	6	5	2	3	6	6	5
AGC	Ser	5	0	2	16	2	6	6	3
AGA	Arg	2	2	4	3	1	8	1	1
AGG	Arg	0	2	0	0	0	0	0	1
GTT	Val	5	6	15	0	7	4	2	3
GTC	Val	8	2	1	11	7	3	3	0
GTA	Val	2	2	5	0	2	4	1	3
GTG	Val	8	4	0	10	6	3	5	3
GCT	Ala	1	3	2	0	8	17	4	2
GCC	Ala	4	7	1	8	6	3	3	3
GCA	Ala	3	7	6	1	4	13	1	1
GCG	Ala	0	1	0	0	0	0	0	0
GAT	Asp	4	5	25	27	3	6	4	2
GAC	Asp	0	2	2	0	1	2	1	3
GAA	Glu	10	6	21	3	6	12	9	6
GAG	Glu	9	5	4	22	5	5	5	5
GGT	Gly	2	1	8	4	0	0	0	0
GGC	Gly	0	0	0	0	0	0	0	0
GGA	Gly	2	1	6	3	1	0	1	0
GGG	Gly	1	0	0	7	1	0	0	0

FIG. 3A

Codon	AA	MSP wt	Edited MSP	MSP wt	Edited MSP	E.coli	Human
TTT	Phe	8	0	0.53	0	0.5	0.35
TTC	Phe	7	15	0.47	1	0.5	0.65
TTA	Leu	25	0	0.66	0	0.11	0.05
TTG	Leu	3	0	0.08	0	0.11	0.09
TCT	Ser	4	1	0.17	0.04	0.27	0.17
TCC	Ser	2	3	0.09	0.13	0.21	0.26
TCA	Ser	10	1	0.43	0.04	0.13	0.11
TCG	Ser	0	0	0	0	0.14	0.07
TAT	Tyr	17	2	0.85	0.1	0.54	0.47
TAC	Tyr	3	18	0.15	0.9	0.46	0.53
TAA	***	0	0				
TAG	***	0	0				
TGT	Cys	10	12	0.83	1	0.45	0.3
TGC	Cys	2	0	0.17	0	0.55	0.7
TGA	***	0	0				
TGG	Trp	0	0	0	0	1	1
CTT	Leu	9	0	0.24	0	0.12	0.11
CTC	Leu	0	0	0	0	0.12	0.22
CTA	Leu	1	0	0.02	0	0.03	0.07
CTG	Leu	0	38	0	1	0.72	0.46
CCT	Pro	4	2	0.4	0.2	0.14	0.24
CCC	Pro	1	6	0.1	0.6	0.11	0.41
CCA	Pro	5	1	0.5	0.1	0.2	0.24
CCG	Pro	0	1	0	0.1	0.54	0.11
CAT	His	3	0	0.75	0	0.64	0.42
CAC	His	1	4	0.25	1	0.36	0.58
CAA	Gln	9	0	1	0	0.31	0.26
CAG	Gln	0	9	0	1	0.69	0.74
CGT	Arg	1	0	0.17	0	0.46	0.09
CGC	Arg	0	0	0	0	0.32	0.19
CGA	Arg	1	0	0.17	0	0.05	0.1
CGG	Arg	0	3	0	0.5	0.06	0.15
ATT	Ile	13	0	0.65	0	0.39	0.23
ATC	Ile	2	20	0.1	1	0.52	0.64
ATA	Ile	5	0	0.25	0	0.08	0.13
ATG	Met	3	3	1	1	1	1
ACT	Thr	3	2	0.19	0.13	0.36	0.2
ACC	Thr	3	13	0.19	0.81	0.38	0.47
ACA	Thr	9	1	0.56	0.06	0.09	0.21
ACG	Thr	1	0	0.06	0	0.17	0.12
AAT	Asn	29	3	0.71	0.07	0.29	0.34
AAC	Asn	12	38	0.29	0.93	0.71	0.66
AAA	Lys	38	0	0.9	0	0.72	0.45
AAG	Lys	4	42	0.1	1	0.28	0.55
AGT	Ser	5	2	0.21	0.09	0.11	0.11
AGC	Ser	2	16	0.09	0.7	0.14	0.29
AGA	Arg	4	3	0.67	0.5	0.08	0.24
AGG	Arg	0	0	0	0	0.03	0.23
GTT	Val	15	0	0.71	0	0.37	0.13
GTC	Val	1	11	0.05	0.52	0.12	0.27
GTA	Val	5	0	0.24	0	0.28	0.09
GTG	Val	0	10	0	0.48	0.23	0.5
GCT	Ala	2	0	0.22	0	0.33	0.31
GCC	Ala	1	8	0.11	0.89	0.18	0.4
GCA	Ala	6	1	0.67	0.11	0.28	0.17
GCG	Ala	0	0	0	0	0.21	0.12
GAT	Asp	25	27	0.93	1	0.48	0.38
GAC	Asp	2	0	0.07	0	0.52	0.62
GAA	Glu	21	3	0.84	0.12	0.67	0.4
GAG	Glu	4	22	0.16	0.88	0.33	0.6
GGT	Gly	8	4	0.57	0.29	0.46	0.15
GGC	Gly	0	0	0	0	0.4	0.44
GGA	Gly	6	3	0.43	0.21	0.06	0.17
GGG	Gly	0	7	0	0.5	0.08	0.24

FIG. 3B

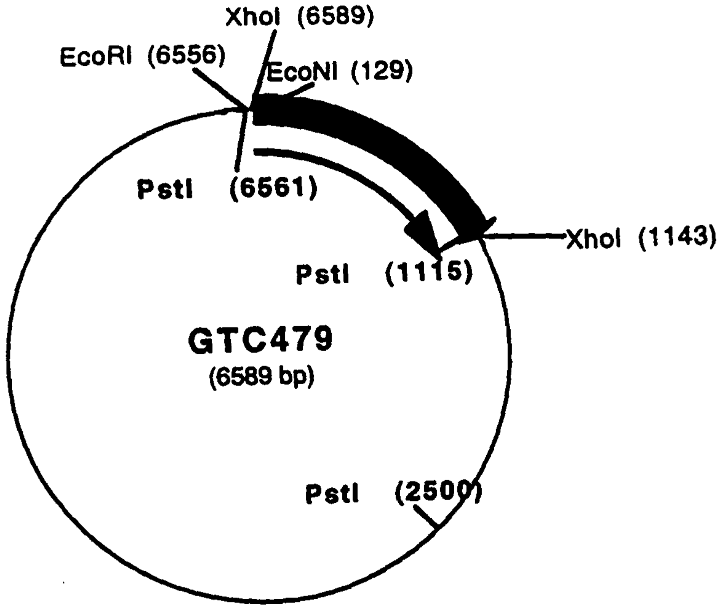

XhoI (6589)
EcoRI (6556)
EcoNI (129)
PstI (6561)
XhoI (1143)
PstI (1115)
GTC479
(6589 bp)
PstI (2500)

FIG. 4A

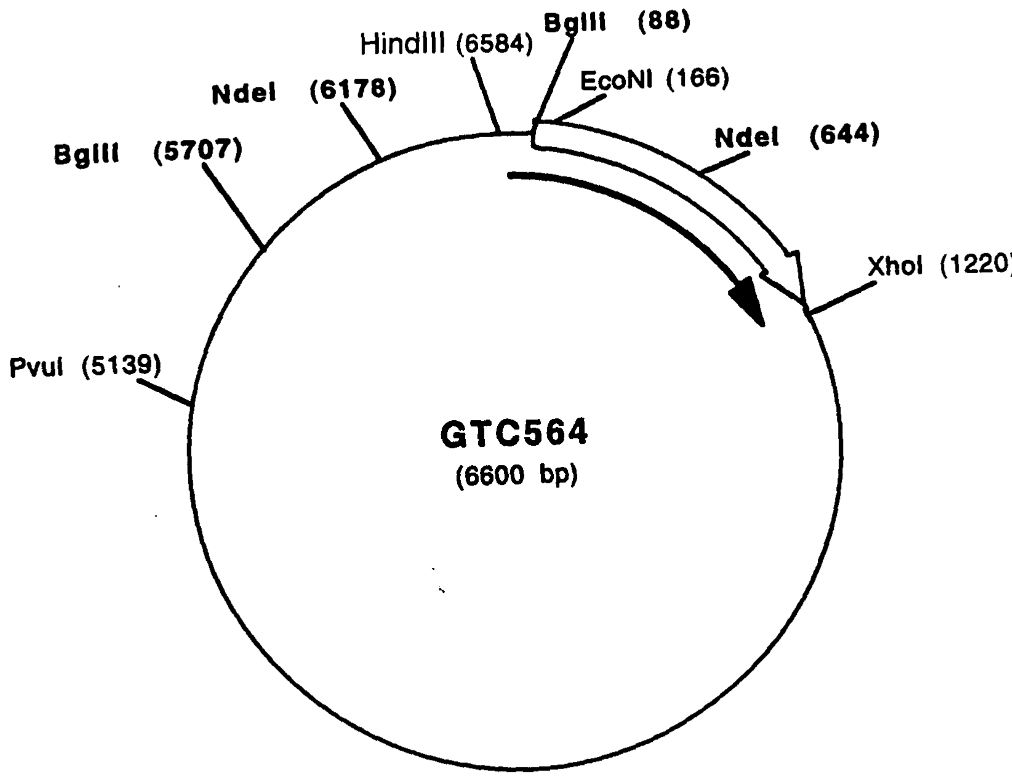

HindIII (6584)
BglII (88)
EcoNI (166)
NdeI (6178)
BglII (5707)
NdeI (644)
XhoI (1220)
PvuI (5139)
GTC564
(6600 bp)

FIG. 4B

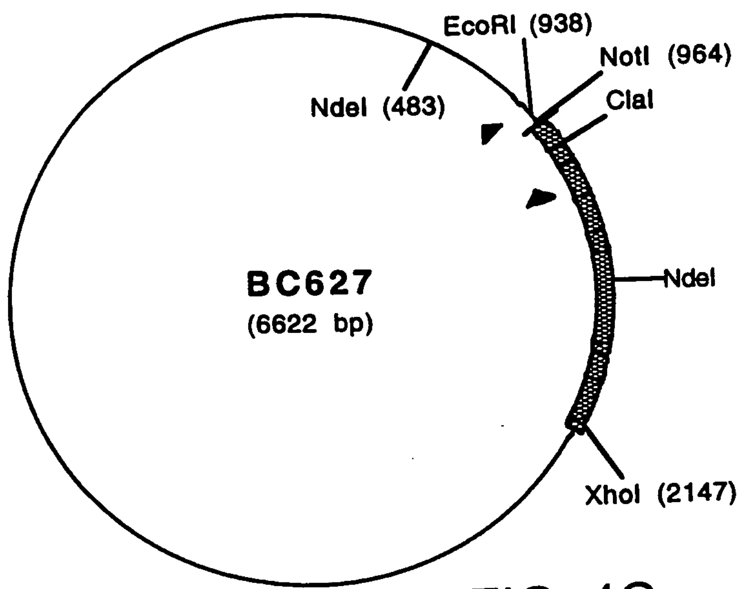


FIG. 4C

Oligos used:

OT1:
TCG ACG AGA GCC ATG AAG GTC CTC ATC CTT GCC TGT CTG GTG GCT CTG GCC ATT GCA AGA GAG CAG GAA GAA CTC AAT GTA GTC GGT A,

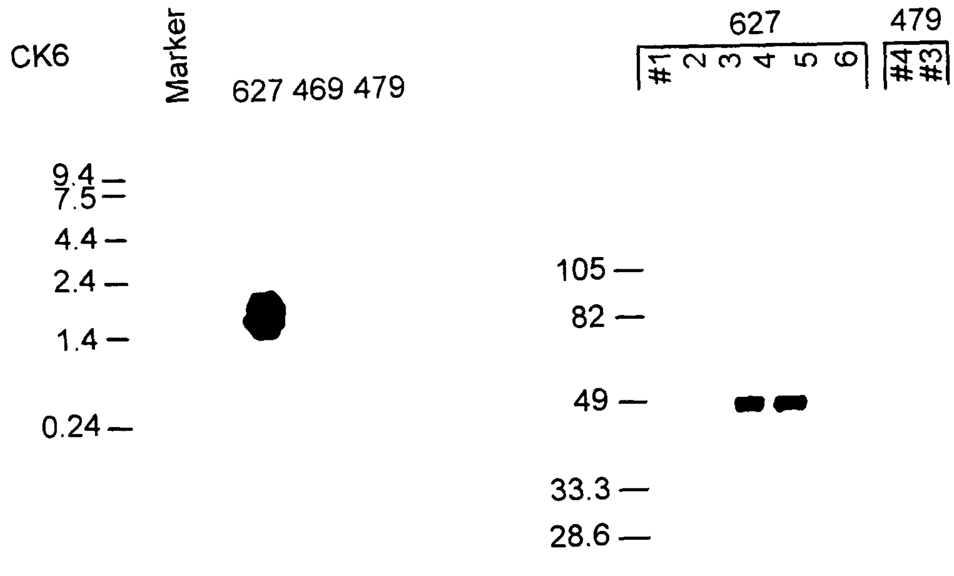
OT2:
GAT CTA CCG ACT ACA TTG AGT TCT TCC TGC TCT CTT GCA ATG GCC AGA GCC ACC AGA CAG GCA AGG ATG AGG ACC TTC ATG GCT CTC G,

MSP1:
AATAGATCTGCAGTAACTCCTTCCGTAATTG,

MSP2:
AATTCTCGAGTTAGTGGTGGTGGTGGTGGTGACTGCAGAAATACCATC

MSP8:
TAACTCGAGCGAACCATGAAGGTCCTCATCCTTGCCTGTCTGGTGGCTCTGG CCATTGCA

FIG. 6



PANEL A

PANEL B

FIG. 5

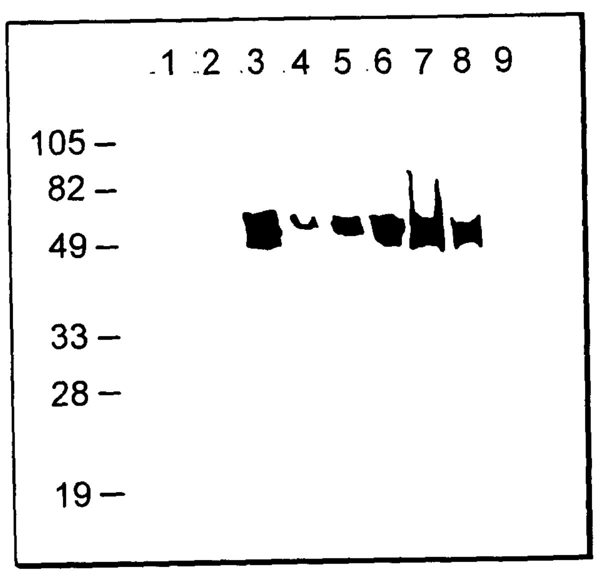


FIG. 10

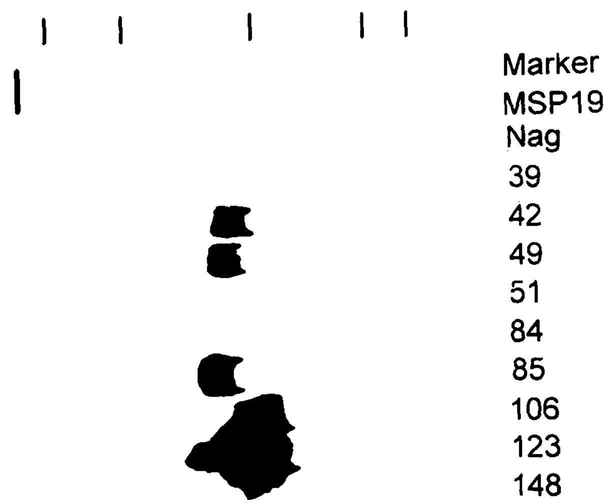


FIG. 13

App[illegible](s): Chen et al.
NO[illegible] MODIFIED MSP-1 NUCLEIC ACID SEQUENCE[illegible]ND METHODS FOR INCREASING MRNA LEVELS AND PROTEIN EXPRESSION IN CELL SYSTEMS

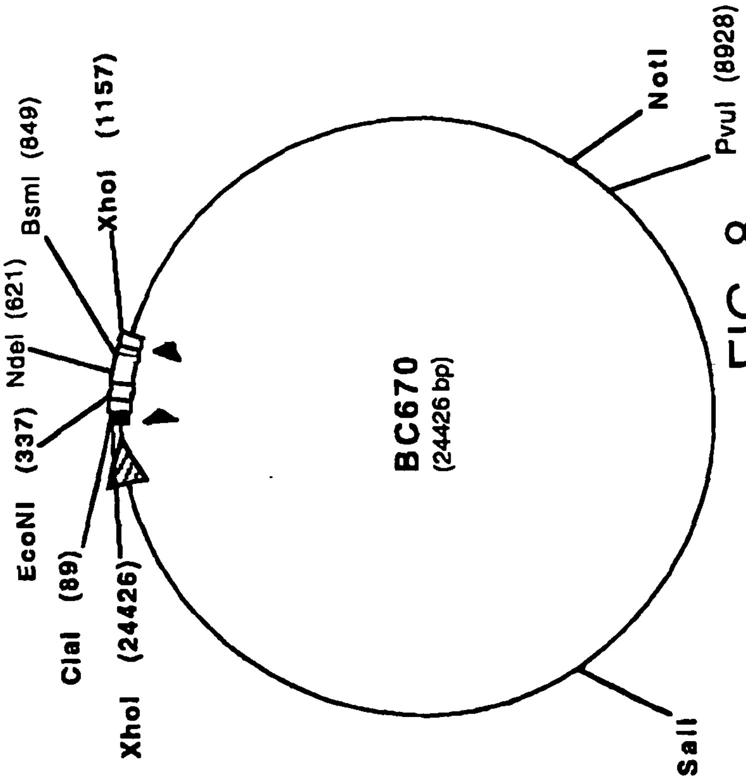


FIG. 8

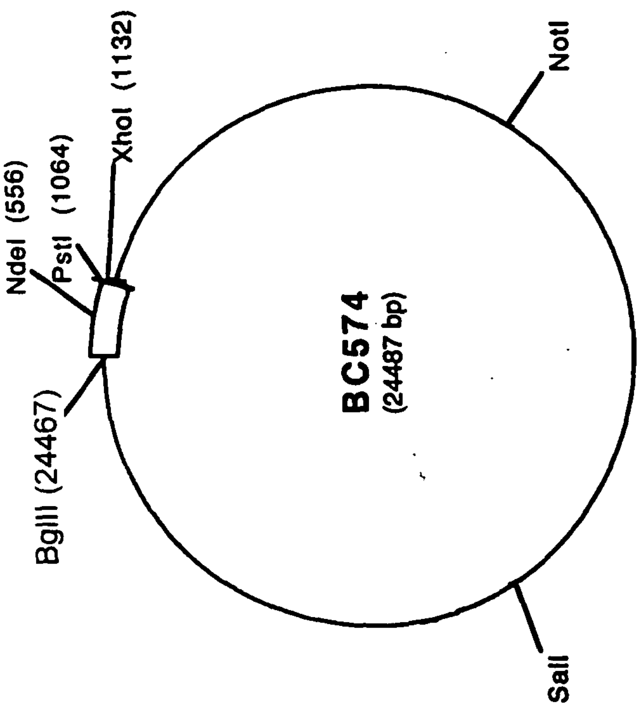


FIG. 7

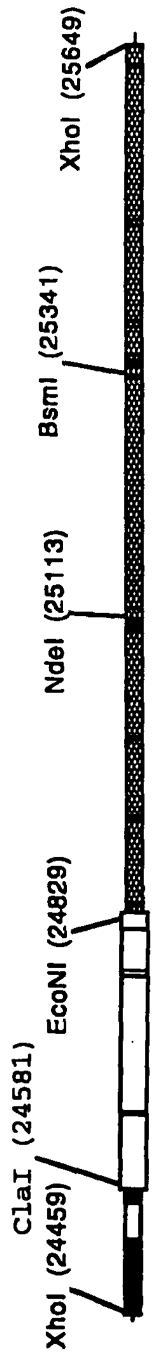


Diagram of BC620

FIG. 9

26 ATGAAGGTCCTCATAATTGCCTGTCTGGTGGCTCTGGCCATTGCAGCCGTCACTCCCTCCGTCATCGATAAC

1▶ M K V L I I A C L V A L A I A A V T P S V I D N
98 ATCCTGTCCAAGATCGAGAACGAGTACGAGGTGCTGTACCTGAAGCCCCTGGCAGGAGTCTACAGGAGCCT

25▶ I L S K I E N E Y E V L Y L K P L A G V Y R S L
169 GAAGAAGCAGCTGGAGAACAACGTGATGACCTTCAACGTGAACGTGAAGGATATCCTGAACAGCAGGTTCAA

48▶ K K Q L E N N V M T F N V N V K D I L N S R F N
241 CAAGAGGGAGAACTTCAAGAACGTGCTGGAGAGCGATCTGATCCCCTACAAGGATCTGACCAGCAGCAACTA

72▶ K R E N F K N V L E S D L I P Y K D L T S S N Y

EcoNI (337)

313 CGTGGTCAAAGATCCCTACAAGTTCCTGAACAAGGAGAAGAGAGATAAGTTCCTGAGCAGTTACAATTACAT

96▶ V V K D P Y K F L N K E K R D K F L S S Y N Y I
385 CAAGGATAGCATTGACACCGATATCAACTTCGCCAACGATGTCCTGGGATACTACAAGATCCTGTCCGAGAA

120▶ K D S I D T D I N F A N D V L G Y Y K I L S E K
457 GTACAAGAGCGATCTGGATAGCATCAAGAAGTACATCAACGATAAGCAGGGAGAGAACGAGAAGTACCTGCC

144▶ Y K S D L D S I K K Y I N D K Q G E N E K Y L P
529 CTTCCTGAACAACATCGAGACCCTGTACAAGACCGTCAACGATAAGATTGATCTGTTCGTGATCCACCTGGA

168▶ F L N N I E T L Y K T V N D K I D L F V I H L E

NdeI (621)

601 GGCCAAGGTCCTGCAGTACACATATGAGAAGAGCAACGTGGAGGTCAAGATCAAGGAGCTGAATTACCTGAA

192▶ A K V L Q Y T Y E K S N V E V K I K E L N Y L K
673 GACCATCCAGGATAAGCTGGCCGATTTCAAGAAGAACAACAACTTCGTCGGAATCGCCGATCTGAGCACCGA

216▶ T I Q D K L A D F K K N N N F V G I A D L S T D
745 TTACAACCACAACAACCTGCTGACCAAGTTCCTGAGCACCGGAATGGTCTTCGAAAACCTGGCCAAGACCGT

240▶ Y N H N N L L T K F L S T G M V F E N L A K T V

BsmI (849)

817 CCTGAGCAACCTGCTGGATGGAAACCTGCAGGGAATGCTGCAGATCAGCCAGCACCAGTGTGTGAAGAAGC

264▶ L S N L L D G N L Q G M L Q I S Q H Q C V K K
888 AGTGTCCCCAGAACAGCGGATGCTTCAGACACCTGGATGAGAGGGAGGAGTGCAAGTGCCTGCTGAACTA

288▶ Q C P Q N S G C F R H L D E R E E C K C L L N Y
958 CAAGCAGGAAGGAGATAAGTGTGTGGAAAACCCCAATCCTACTTGTAACGAGAACAATGGAGGATGCGATG

311▶ K Q E G D K C V E N P N P T C N E N N G G C D
1029 CCGATGCCAAGTGTACCGAGGAGGATTCAGGAAGCAACGGAAAGAAGATCACCTGCGAGTGTACCAAGCCT

335▶ A D A K C T E E D S G S N G K K I T C E C T K P

XhoI (1157)

1100 GATTCTTATCCACTGTTCGATGGtATtTTCTGCAGTCACCACCACCACCACCACTAACTCGAGGAT

359▶ D S Y P L F D G I F C S H H H H H H • L E D

FIG. 11

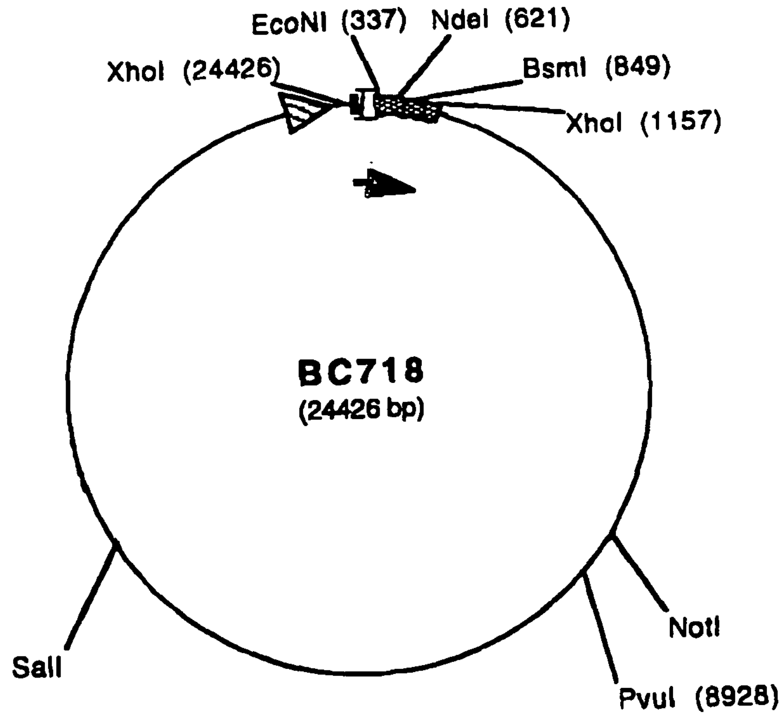

EcoNI (337)
NdeI (621)
XhoI (24426)
BsmI (849)
XhoI (1157)
BC718
(24426 bp)
SalI
NotI
PvuI (8928)

FIG. 12